The Astrophysical Journal, 954:87 (15pp), 2023 September 1 https://doi.org/10.3847/1538-4357/acdd7b

Comparative Analysis of TRGBs (CATs) from Unsupervised, Multi-halo-field Measurements: Contrast is Key

J. Wu[1,2], D. Scolnic[2], A. G. Riess[3,4], G. S. Anand[3], R. Beaton[3,4,5,6], S. Casertano[3], X. Ke[2], and S. Li[4]
[1] Kuang Yaming Honors School, Nanjing University, Nanjing, Jiangsu 210023, People's Republic of China; jiaxiwu1018@gmail.com
[2] Department of Physics, Duke University, Durham, NC 27708, USA; daniel.scolnic@duke.edu
[3] Space Telescope Science Institute, Baltimore, MD 21218, USA
[4] Department of Physics and Astronomy, Johns Hopkins University, Baltimore, MD 21218, USA
[5] Department of Astrophysical Sciences, Princeton University, 4 Ivy Lane, Princeton, NJ 08544, USA
[6] The Observatories of the Carnegie Institution for Science, 813 Santa Barbara Street, Pasadena, CA 91101, USA
Received 2022 November 13; revised 2023 May 23; accepted 2023 May 24; published 2023 August 24

Abstract

The tip of the red giant branch (TRGB) is an apparent discontinuity of the luminosity function (LF) due to the end of the red giant evolutionary phase and is used to measure distances in the local universe. In practice, tip localization via edge detection response (EDR) relies on several methods applied on a case-by-case basis. It is hard to evaluate how individual choices affect a distance estimation using only a single host field while also avoiding confirmation bias. To devise a standardized approach, we compare *unsupervised*, algorithmic analyses of the TRGB in *multiple* halo fields per galaxy. We first optimize methods for the lowest field-to-field dispersion, including spatial filtering, smoothing, and weighting of LF, color band selection, and tip selection based on the number of likely RGB stars and the ratio of stars below versus above the tip (R). We find R, which we call the tip *contrast*, to be the *most important* indicator of the quality of EDR measurements; higher R selection can decrease field-to-field dispersion. Further, since R is found to correlate with the age or metallicity of the stellar population based on theoretical modeling, it might result in a displacement of the detected tip magnitude. We find a *tip-contrast relation* with a slope of -0.023 ± 0.0046 mag/ratio, an $\sim 5\sigma$ result that can be used to correct these variations in the detections. When using TRGB to establish a distance ladder, consistent TRGB standardization using tip-contrast relation across rungs is vital to make robust cosmological measurements.

Unified Astronomy Thesaurus concepts: Distance indicators (394); Stellar distance (1595); Standard candles (1563)

1. Introduction

The tip of the red giant branch (TRGB), the location of a prominent break in the luminosity function (LF) of red giant stars due to the sudden onset of core-helium burning near an absolute magnitude of $M_I \sim -4$ mag, provides one of the few primary distance indicators useful to distances of $D > 10$ Mpc, thus making it an important tool for measuring the rate of the expansion of the universe (Lee et al. 1993; Serenelli et al. 2017). It has been relied upon heavily to measure distances to nearby galaxies for determining precise distance-dependent parameters (McQuinn et al. 2014; Crnojević et al. 2019; Shen et al. 2021) and discerning galaxy flow patterns (Anand et al. 2019; Shaya et al. 2020; Tully et al. 2022). Recently, it has also been employed as a centerpiece of some distance ladders used to measure the Hubble constant (H_0) (Jang & Lee 2017; Freedman et al. 2020; Blakeslee et al. 2021; Anand et al. 2022; Dhawan et al. 2022).

The tip is often identified using a peak or local maximum of a derivative function or Sobel filter applied to a histogram of stars ordered by magnitude (Lee et al. 1993). A challenge for evaluating the quality of such a measurement is that it provides only a single number for a single field and without a goodness-of-fit measure, thus offering no means of testing that are independent of the application. This is different than multi-source distance indicators like Cepheid or Mira variable stars for which multiple, comparable examples per host are available or other indicators like Type Ia supernovae (SNe Ia) with multiple SNe Ia *siblings* seen in the same host galaxy (Scolnic et al. 2020). Because the apparent edge of the LF may be contaminated by, e.g., young asymptotic giant branch (AGB) stars or may be confused with stochasticity elsewhere in the LF, the measurements are often performed in a *supervised* fashion wherein additional techniques are used to select the *most likely* TRGB among multiple peaks in the EDR.

The number of techniques used to pick the tip are multifold and include the general framework of tip detection (edge detection, e.g., Freedman et al. 2019, or fitting a model LF, e.g., Anand et al. 2022), filtering the data to remove young populations (spatially, e.g., Anand et al. 2018 and Jang et al. 2018; or by color selection e.g., Jang et al. 2021), the degree of smoothing applied to the LF (Hatt et al. 2018; Beaton et al. 2019), and which quality cuts to employ (e.g., the ratio of RGB to AGB stars, Hoyt et al. 2021; the ranking of tips, Hoyt 2021; the number of stars below the peak, Madore & Freedman 1995; or projected distance into the halo, Jang et al. 2021). Further complicating the analysis is the spatial inhomogeneity of many fields and their stellar populations leading to a qualitative selection of fields as either suitable or not when quantitative metrics applied locally in the field are warranted, such as illustrated in Radburn-Smith et al. (2011). As a result, different options, parameters for these options, or combinations of methods are often used for different galaxies. For instance, some papers applied a color selection (Jang et al. 2021) while some did not (Hoyt et al. 2019), and Hatt et al. (2018) applied two different smoothing scales to NGC 1316 and NGC 1448.

There are two areas in which measurement inconsistencies could lead to systematic errors along a distance ladder. First, tip measurements must be performed with a self-consistent algorithm to avoid bias, estimate uncertainty introduced by variants of the analysis, and confirmation bias (being insensitive to where the tip is expected to be from SNe Ia measurements or others). The second challenge arises from astrophysical variance. The TRGB refers to a sudden break in the magnitude histogram, whose physical interpretation is red giant stars experiencing rapid luminosity changes after the He flash (Serenelli et al. 2017). For measuring H_0, the measurements have typically been in the stellar halo where populations are assumed to be simple: old and metal-poor. However, halo sight lines contain populations of stars with uneven star formation histories and varying metallicities (McQuinn et al. 2019; Cohen et al. 2020). Even with minimal measurement noise and in the relatively metallicity insensitive *I* band, the tip may be seen to be *fuzzy* due to the presence of AGB stars, which can be brighter than the tip and of similar color (Conn et al. 2012; Beaton et al. 2019; Mager & Hoyt 2020). It is therefore essential to standardize the tip measure across the distance ladder in case there is either a bias in the tip location due to the loss of tip contrast or an intrinsic spread in the tip luminosity that depends on the specific population producing the tip.

To optimize analysis choices, quantify the corresponding uncertainties and variance of TRGB measurements, and determine the fraction of fields that can have robust measurements, we leverage nearby GHOSTS data (Galaxy Halos Outer Disks Substructure Thick Disks Star Clusters; Radburn-Smith et al. 2011) where multiple fields per host, up to 11 fields, have been observed with a high signal-to-noise ratio (S/N). By comparing the tip detection across halo fields in the same host, we can ascertain the efficacy of different analysis choices. The structure of this paper is as follows. In Section 2, we describe the data and what is provided. In Section 3, we present a compilation of analysis options. In Section 4, we determine the consistency of tip detection given various choices. In Section 5, we discuss how the results of this analysis can be extrapolated to its usage in the distance ladder.

2. Data

For this analysis, we make use of the GHOSTS survey (Radburn-Smith et al. 2011), which observed resolved stellar populations of 14 disk galaxies within 17 Mpc, and the outer disks and halos were imaged with the Hubble Space Telescope (HST) Advanced Camera for Surveys (ACS). The sample was observed with filters ACS F606W (similar to the *V* band) and F814W (similar to the *I* band), which are commonly used for TRGB measurements (Freedman et al. 2019; Anand et al. 2021a). In this paper, we utilize 10 disk galaxies from the GHOSTS survey with multiple halo fields and discard other galaxies that only contain no more than one halo field. The GHOSTS photometry we employ was measured using the `DOLPHOT` package (Dolphin 2000) with photometric uncertainties derived from artificial stars tests and used source size and shape to remove contaminants (e.g., background galaxies). Because many of the GHOSTS hosts are closer (e.g., M81 at 3 Mpc) than SN Ia hosts used for distance ladder measurements (∼20 Mpc), the S/N of the data is greater thus providing better resolution of underlying astrophysical variance of the tip. The GHOSTS team has posted their photometry files for multiple fields of local galaxies on their website.[7]

In the GHOSTS program, fields were chosen along the major and minor axes of each galaxy, principally to probe the halo structure. The precise field locations were chosen to avoid bright foreground stars, to allow HST to use all available roll angles, and sometimes to sample features such as the disk truncation or previously identified stellar streams. The size of each field is roughly 11.1 arcmin^2, a single HST+ACS pointing. The depth of the survey at 50% completeness is 2.7 mag below the tip of the red giant branch (TRGB).

The number of fields observed per galaxy ranges from one to eight and in some cases with an additional up to eight observations contributed from the HST archive. Although the archival observations come from a variety of programs and do not necessarily have the same depth as the ones from the GHOSTS survey, they are merged in the data set provided on the GHOSTS webpage and we take those directly from GHOSTS. The archival data are typically $1 \sim 1.5$ mag fainter than that of the GHOSTS survey, though are processed with the same photometry and calibration packages as the images taken as part of the GHOSTS survey. There are also extreme circumstances where archival data are $2 \sim 2.5$ mag fainter, such as NGC 3031 Halo Field 01 and Halo Field 11, see Figure 2. A graphic of the locations of GHOSTS fields is shown for NGC 3031 in Figure 1. There is a range of distances from the center of the galaxy (disk field) up to 30 kpc away (e.g., Halo Fields 03, 06, and 11). We limited the fields to those where TRGB is most effective; that is, we excluded fields that visually were predominantly inside the disk (25th visual mag isophote). We retained fields with a small fraction inside the disk, which helped facilitate the development and demonstration of an algorithm that can empirically find and use the disk boundary rather than the use of more subjective or nonuniform criteria for identifying *halo* and *disk*.

Examples of the color–magnitude diagrams (CMDs) as provided by this analysis are shown in Figure 2. The diversity of CMDs is apparent by eye; for some, the main sequence (MS) is visible, and for some, the slope of the RGB sequence appears to change. Radburn-Smith et al. (2011) also measured distances to multiple fields within each galaxy in the GHOSTS survey. Radburn-Smith et al. (2011) only report distances for fields away from the disk and where the fit converges past its own internal quality standards and with some deference to the sample mean, e.g., Halo Field 02-06 shown in Figure 1 and Figure 2. Further, the analysis by Radburn-Smith et al. (2011) was *supervised* in the sense that the measurement procedure was customized to each field rather than the algorithmic approach we attempt here, varying a number of analysis choices for global optimization. The goal of our analyses is rather than determining the best distances to individual galaxies, to measure the nature of the variation of TRGB measurements across multiple fields of the same galaxy.

3. Methodology

In this section, we develop an algorithm for the automatic detection of candidate TRGB without a priori knowledge of the distance. We show a descriptive graphic of the multiple steps and measurements needed in Figure 3. We summarize the analysis choices in Table 1. We focus on methods that can be

[7] https://archive.stsci.edu/prepds/ghosts/

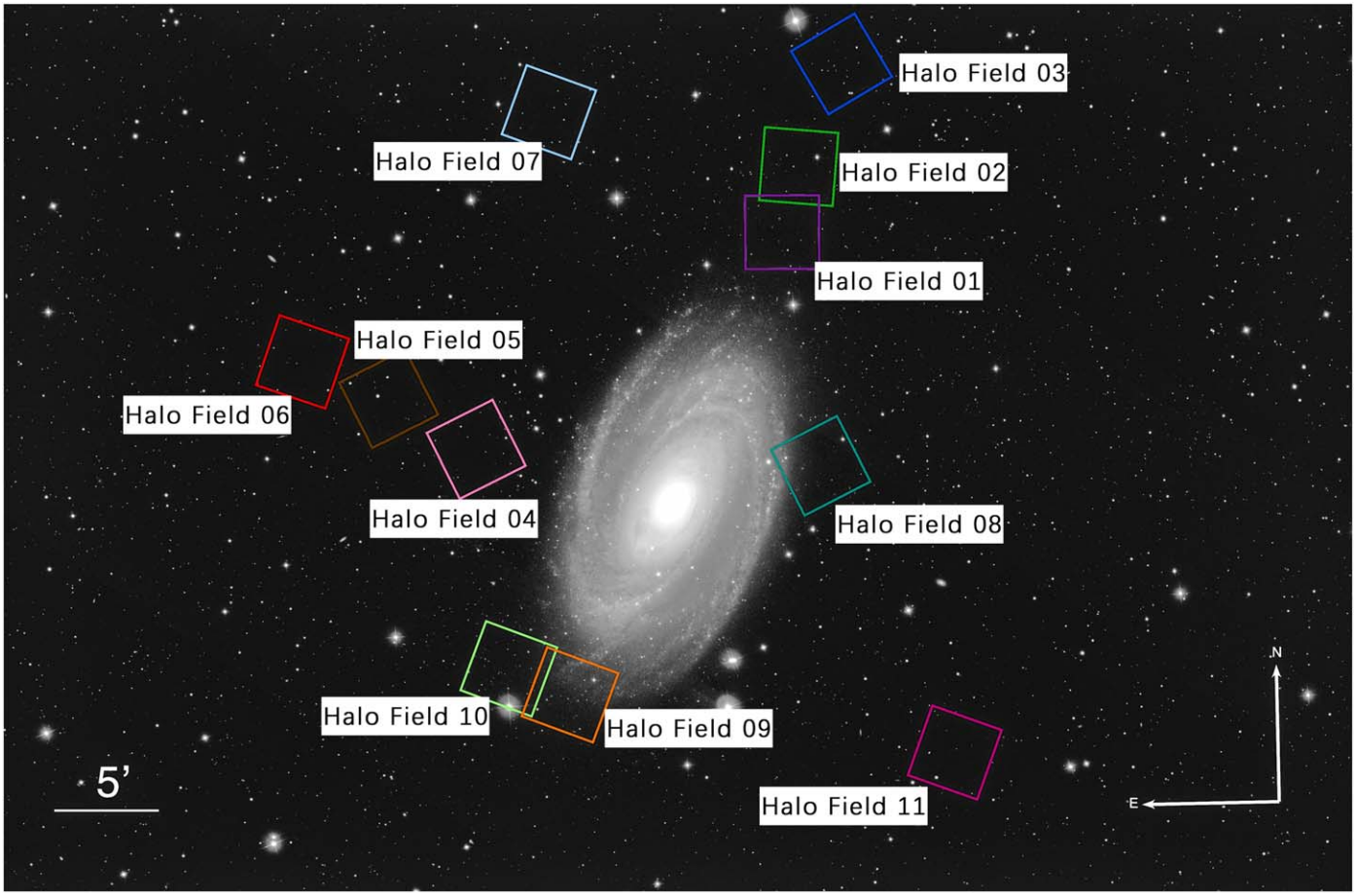

Figure 1. A footprint of NGC 3031 produced from SDSS *gri* data (Alam et al. 2015), generated with the online Imaging Mosaics tools (https://dr12.sdss.org/mosaics/). The locations of the fields observed as part of the GHOSTS program are indicated and labeled. The GHOSTS fields are each 1 pointing of HST+ACS corresponding to 11.1 arcmin^2. All halo fields are located at or past a 25 mag arcsec^{-2} isophot.

most consistently applied using the measured CMD. In order to develop a uniform analysis procedure that is widely applicable, we do not use ancillary information, such as H I maps (e.g., at $D < 10$ Mpc, see Walter et al. 2008; Heald et al. 2011), which may not be always available for galaxies in the distance ladder or will lack the necessary depth and resolution (e.g., for SN hosts at $D \sim 20$ Mpc, see Condon 1987; Condon et al. 1996). Although it may be possible to produce a *better* measurement of the EDR in an individual field through customization, the value of the unsupervised approach we seek here is that it is uniform, readily reproduced, and less sensitive to some types of bias. We will show in Section 4 that we can reach a low internal dispersion (<0.05 mag) for a range of fields, thus demonstrating that our approach is viable.

The rest of this section presents a comprehensive description of each of the choices in our analysis procedure.

3.1. Spatial Clipping of Young Populations

Jang et al. (2021) and Anand et al. (2022) show that younger stellar populations can contaminate the tip, primarily through the presence of intermediate-age AGB stars and also from supergiants blueward of the RGB. One way to eliminate the contamination is using the spatial distribution of MS stars as a proxy for regions with younger stellar populations. Alternate routes would be a cut on the distance from the disk of the galaxy, which depends significantly on the shape of the galaxy, or a cut on the local sky brightness (Anand et al. 2018). Here we follow the conceptual approach of Anand et al. (2022), i.e., use MS stars to identify the position of young stellar populations, and develop a specific procedure for this step.

We illustrate our algorithm in Figure 4. First, we correct the CMD for Milky Way (MW) reddening (Schlafly & Finkbeiner 2011). As shown in Figure 2, there is a vertical red line, which equals F606W−F814W = 0.3 mag after MW extinction correction, as well as a horizontal green line, which is the magnitude where S/N reaches 10. Then we utilize the MS stars bluer than the vertical red line and brighter than the horizontal green line as a proxy for Population I stars, as marked blue in the bottom-left panel of Figure 4, and create a density map showing the number of blue stars per 16 arcsec^2 area. The green line is necessary here to prevent counting in parts of the CMD where the MS and RGB merge into each other. We also calculate the number of other stars (black dots in the bottom-left panel of Figure 4), as well as the ratio of these two densities. Density plots of the ratio and the blue star population are shown in the first two panels of Figure 4. We smooth the density map of MS stars using a two-dimensional Gaussian-windowed, Locally Weighted Scatterplot Smoothing (GLOESS) algorithm, which uses a two-dimensional Gaussian weighting function as

$$w([i, j] \rightarrow [m, n]) = e^{-\frac{(x(i)-x(m))^2+(y(j)-y(n))^2}{2\sigma^2}}, \quad (1)$$

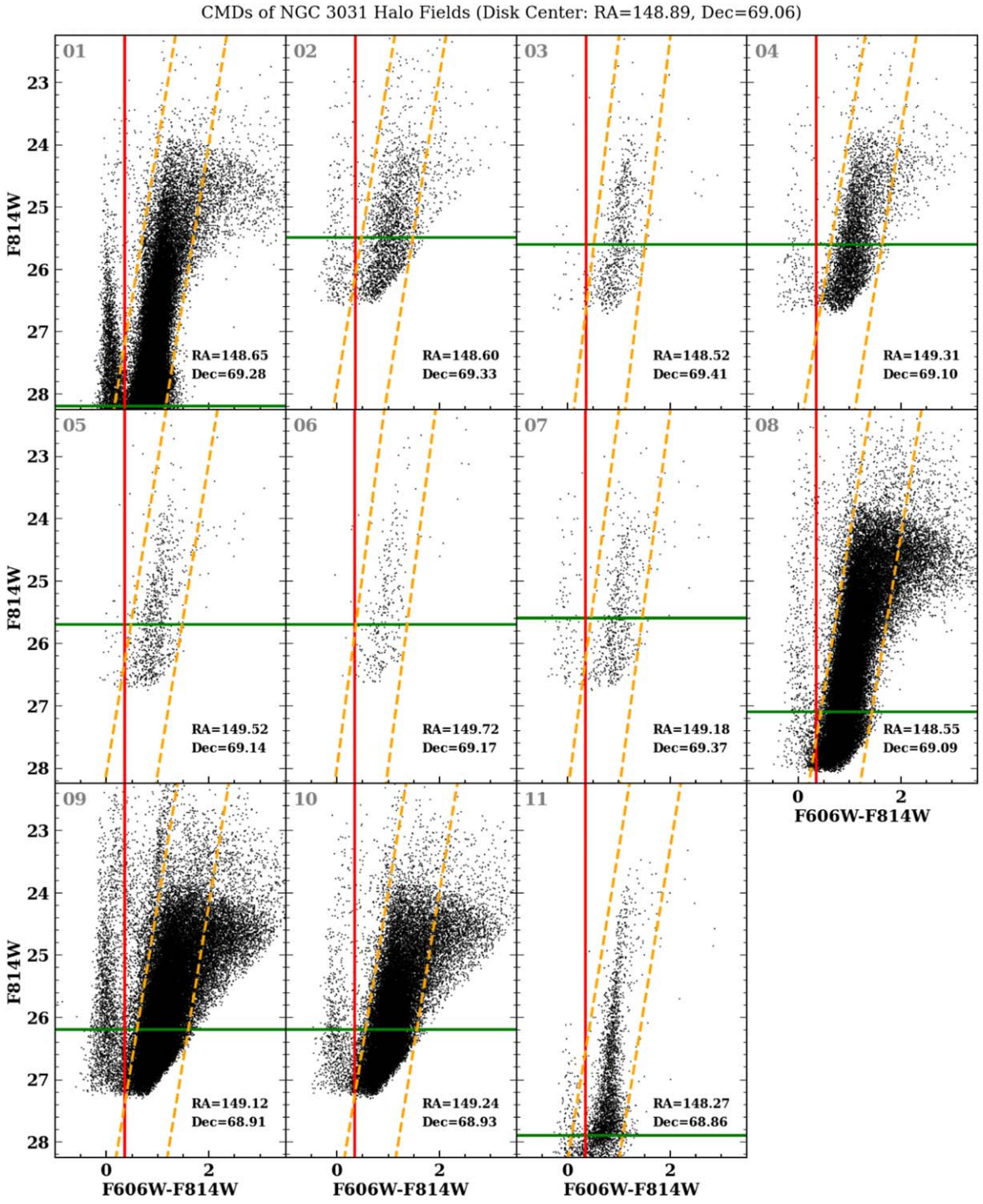

Figure 2. CMDs of individual fields in NGC 3031 with labels corresponding to the fields presented in Figure 1. The field numbers are labeled at the top left corner of each panel and match those in Figure 1, and the R.A. and decl. of each field are labeled at the bottom-right corner, in units of degree in epoch J2000. Halo Fields 01 and 08-11 are from HST external, archival data that has been involved in the analyses of the GHOSTS survey, and data of all 11 halo fields are from the GHOSTS compilation. The properties of the CMD vary from field to field; some of this can be attributed to contamination from younger stellar populations where the disk and halo overlap. This study aims to quantify the remaining differences and their impact on TRGB measurements. The red vertical line represents our separation of the MS stars from evolved stars, like the AGB and RGB, and the green line shows where the S/N reaches 10. Both lines are used in the spatial clipping algorithm, as discussed in Section 3.1. The orange lines denote the optimized color that maximizes the number of stars included on the RGB branch. The color and magnitude selections are discussed in Section 3.

where functions x and y are the bin center in units of pixels. The smoothing scale is set by the σ parameter in Equation (1). For this analysis, we fix the smoothing scale for all fields, independent of distance. For this algorithm, we must determine the number for the smoothing scale as well as the spatial quality cut level and manner. Using the metrics discussed in Section 4.1, we find the optimal manner to cut is using the number density of blue MS stars rather than the ratio of blue-to-red stars. We use a cut level that is relative to the peak number density, and we determine from the metrics that this value

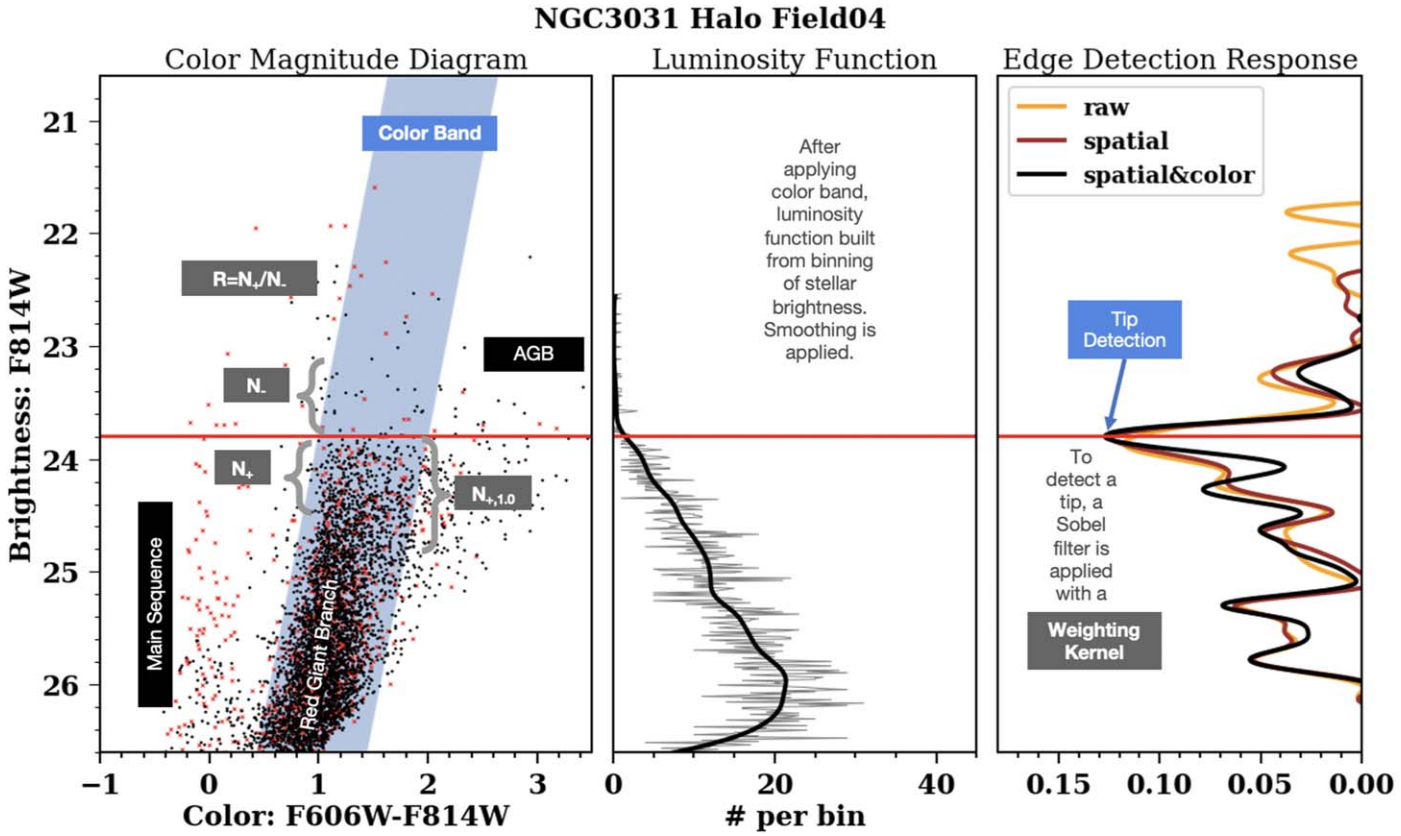

Figure 3. A demonstration of how the TRGB magnitude is determined from the CMD for one of the halo fields in NGC 3031 (#4). For each step, analysis choices must be made including color cuts and spatial clipping (panel 1), and smoothing/weighting of the LF (panels 2 and 3). The red line crossing all three panels is the tip detection line and matches the tallest peak of the black curve in panel 3. The red "x" points in panel 1 indicate the stars cut using spatial clipping as discussed in Section 3.1. The N_- and N_+, as discussed in Section 3.4, indicate the range of magnitudes 0.5 mag brighter and fainter than the red TRGB magnitude, respectively. The gray distribution in panel 2 represents the LF and the bold black curve the smoothed LF. In panel 3, the *raw* EDR is with no color band or spatial clipping applied; the *spatial* EDR is with only a spatial clipping applied; the "spatial&color" EDR is with both a color band and spatial clipping applied.

Table 1
Different Analysis Parameters for Determining the Most Likely TRGB from a CMD

Analysis Parameters		Range Considered			Section
Spatial clipping		Clipped	Not clipped		3.1
Color band	Slope	<0			3.2
		∞ (Vertical)			
	Intercept	variable			
	Width	$0.5 \sim 1.5$			
Smoothing factor		$0.03 \sim 0.12$			3.3
Weighting kernel		Simple	Poisson	Hatt	
$R = N_+/N_-$		$\geqslant 1.5$			3.4
$N_{+,1.0}$(1 mag below tip)		$\geqslant 50$			

Notes. On the left, the different options are given, and in the middle column, the different values for each choice are given. These are explained in detail in Section 3, with subsection links displayed in the right column. Our goal is to emulate the combination of techniques used in the literature on TRGB edge detection rather than devise new methods.

should be $\sim$10%, such that all areas with a number density of blue stars $>$10% of the peak number density for that field are removed. Furthermore, we find a smoothing scale σ of 160 pixels is best; we show the results of smoothing in the top right of Figure 4. We present the success of this algorithm in the bottom row of Figure 4. The bottom-left and right panels show the original and clipped CMDs, respectively, while the bottom middle panel shows how the field is carved out. The figure shows that the younger population as seen from the CMD (seen in the bottom-left panel) is removed to a much greater extent than the RGB (as seen in the bottom-right panel).

3.2. Color Cuts

To further reduce contamination of the TRGB it is common to apply a slanted band in color–magnitude space that follows the RGB branch (Hatt et al. 2018; Jang et al. 2018), so the measurement is limited to the bluer, metal-poor region that is found to have a nearly constant absolute magnitude with color. An example of such a band is shown in Figure 3. A challenge of defining a proper band is that the mean color or slope of this branch may vary across fields or galaxies and that there is a balance to be struck between using a narrow range of color for population uniformity and retaining more stars for better statistics. In the literature, the type of color band has varied substantially. For example, Hatt et al. (2017, 2018) used a band to include the largest number of stars along the RGB branch, typically of width 0.5–2.0 mag. In Freedman et al. (2019), it is said that no band is used for the distance ladder measurement. In some analyses, when measured in F814W and F606W, the slope of the RGB branch is set to a fixed parameter (e.g., $k = -6$ mag $\cdot$ color^{-1} in Jang et al. 2018). In other analyses like that of Hoyt et al. (2019, 2021), there is no color band applied (which we call *vertical* as it is equivalent to a band slope of ∞).

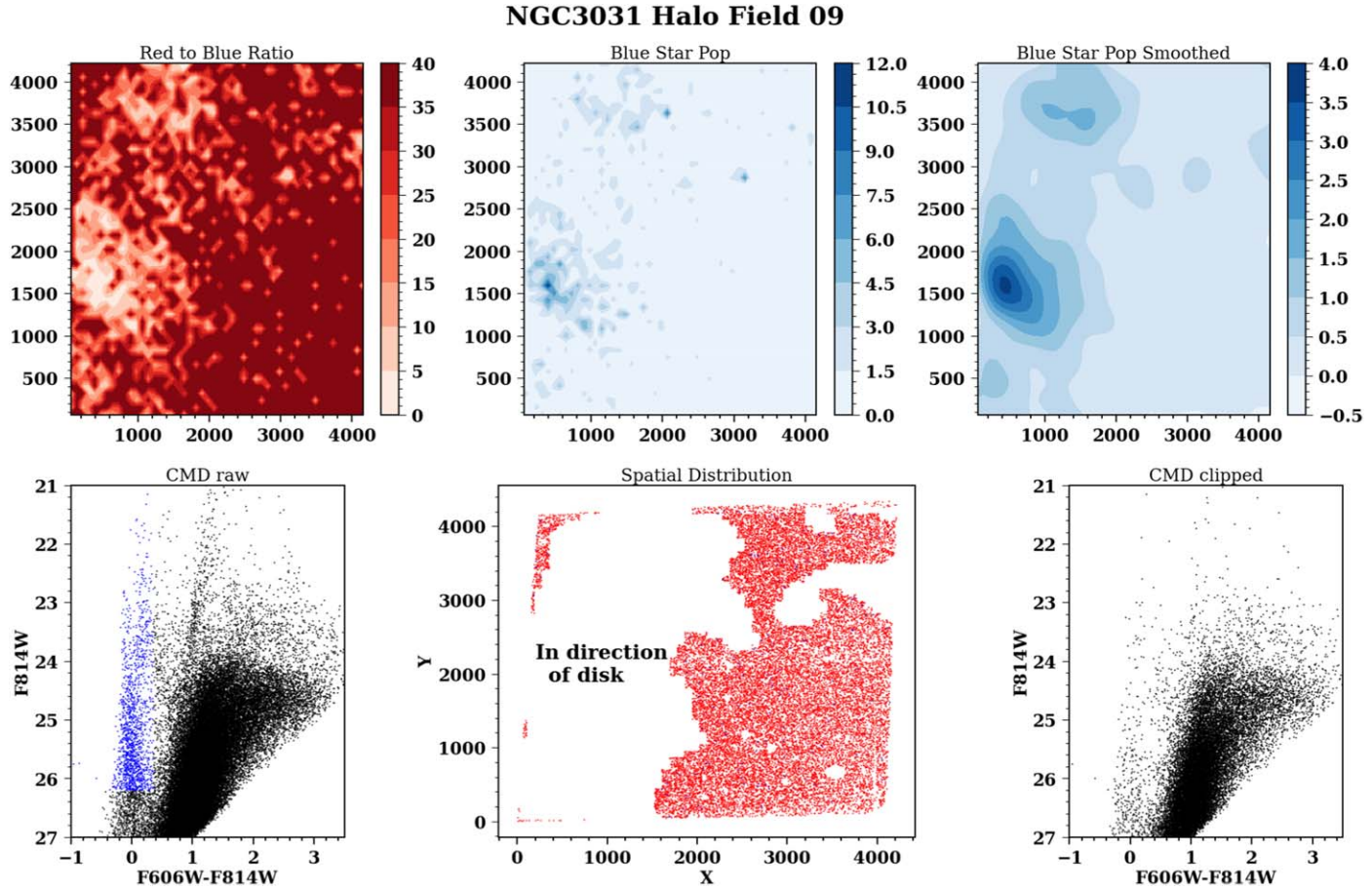

Figure 4. A demonstration of the spatial clipping used for this analysis to suppress contamination from young stars and AGB stars. As shown in Figure 2, we use a red vertical line and a green horizontal line to find the location of the MS, and we calculate the number density of *blue stars* (i.e., stars bluer than the red line and brighter than the green line in color–magnitude space) and *red stars* (i.e., all other stars on the CMD) as well as the ratio of red stars to blue stars. The number densities and ratio are calculated with a resolution size of 16 arcsec^2. We plot the ratio and blue star population in the top row (panel 1 and panel 2, respectively) across a spatial region of 11.1 arcmin^2, which is the size of an entire field in the GHOSTS survey. The units of the panels in the first row are ACS pixels. In panel 3, we perform a 2D-GLOESS smoothing on the blue star population, and we find through optimization that cutting areas based on the smoothed number density of blue stars yields the best tip measurements, as quantified in Section 4. In the bottom row, we show in the left panel the original CMD of this field, with blue stars marked blue. The middle panel shows how this space is carved out using a 2D-GLOESS filter, the white empty space is the region clipped, and the red and blue dots elsewhere are the remaining red stars and blue stars. We present a tag of "in direction of disk" in the space showing the region near the galaxy disk, which demonstrates the efficiency of our spatial clipping algorithm. In the bottom-right panel, we show how the blue main-sequence cluster in the CMD, together with non-MS young star populations such as supergiants, is heavily suppressed after this spatial clipping.

We parameterize a general color band with two boundaries, a blue and red limit with

$$\begin{aligned} \mathrm{F814W_{blue}} &= s * (\mathrm{F606W} - \mathrm{F814W}) + b \\ \mathrm{F814W_{red}} &= s * (\mathrm{F606W} - \mathrm{F814W} - w) + b, \end{aligned} \quad (2)$$

where the slope is s (Δ mag/Δ color) and the width of the band is w and intercept b. For an unsupervised algorithm, we fix the bandwidth w, and provide a range of the slope s ($-7 \sim -5$ for F606W/F814W CMDs, and $-3 \sim -1$ for F555W/F814W CMDs), and the values for s and b are found to maximize the number of stars in the band. This optimization can be done iteratively with the more global optimization of the width and slope range (see the Appendix) and redetermining the offset.

3.3. Sobel Detection

We utilize the ACS F606W and F814W photometry from the GHOSTS data set to construct CMDs. To create the LFs, we bin the stars by their ACS F814W magnitude in 0.01 mag intervals, and use the GLOESS method to smooth the LF (Persson et al. 2004). GLOESS distinguishes itself from other local regression smoothing techniques by applying weight from all bins with an emphasis locally such that the ith bin when smoothing locally at the jth bin as

$$w(i \rightarrow j) = e^{-\frac{(\mathrm{mag}(i)-\mathrm{mag}(j))^2}{2\sigma^2}}, \quad (3)$$

where mag(i) is defined at the bin center and σ is the smoothing scale. Hatt et al. (2018) offer an algorithmic approach to the optimal smoothing method through the use of artificial star tests and find that the smoothing scale should roughly equal the level of the photometric uncertainty near the tip magnitude. However, this approach does not appear to be adopted for other TRGB measurements (e.g., Freedman et al. 2019). We vary the smoothing scale as an option, as shown in Table 1 and discussed in the Appendix, to minimize the field-to-field dispersion of the tip measurement around the mean for each galaxy.

The second step is generating the edge detection response (EDR) curve by applying a Sobel detection. The Sobel kernel

[−1,0,+1] was first introduced into the context of TRGB by Lee et al. (1993) and is equivalent to a centered first derivative of the function

$$\mathrm{EDR}(i) = \mathrm{LF}(i+1) - \mathrm{LF}(i-1), \quad (4)$$

so it gives the maximum response at the position where LF changes fastest in star number density, that is, the magnitude of TRGB. This corresponds to the physical interpretation of TRGB, due to a termination of the RGB branch, thus resulting in a sudden drop of the stellar population. In Table 1, we denote this first-derivative Sobel detection as *simple*. One modification of this method considers the S/N by applying a Poisson weighting

$$\mathrm{EDR}(i) = \sqrt{\mathrm{LF}(i+1)} - \sqrt{\mathrm{LF}(i-1)}, \quad (5)$$

which we denote as *Poisson*. Hatt et al. (2017) employed another form of Poisson weighting

$$\mathrm{EDR}(i) = \frac{\mathrm{LF}(i+1) - \mathrm{LF}(i-1)}{\sqrt{\mathrm{LF}(i+1) + \mathrm{LF}(i-1)}}, \quad (6)$$

which we denote as *Hatt*. Since the model of LF predicts an exponential growth of star population below the tip (Méndez et al. 2002), the *simple* Sobel filter might result in false detection in the faint end. The Poisson weighting can suppress faint, false peaks. For our analysis, we use Equation (6) for the Sobel detection weighting, and the optimization process is shown in the Appendix.

3.4. Tip Detection

After generating the EDR curve under a specific smoothing and weighting technique, we introduce a tip detection method that ultimately identifies the strongest tip as well as tips of comparable size. We first find all local maximum points that satisfy

$$\mathrm{EDR}(i-1) < \mathrm{EDR}(i) > \mathrm{EDR}(i+1) \quad (7)$$

and denote the set of these points as $[\mathrm{EDR_{tot}}]$.

We define two properties of candidate peaks measured after applying a spatial and color cut: the contrast and the number of stars below the tip. These can be used as both quality indicators for tip measurements and used for further correlation studies.

1. *Contrast, R*: For every F814W magnitude in the observable range, we define a contrast value that is the ratio of the number of stars below versus above that magnitude, given bounds of 0.5 mag. This can be defined as

$$R = N_+/N_-, \quad (8)$$

where N_+ and N_- are the number of stars between the given magnitude and + or −0.5 mag respectively. This ratio resembles that used by Hoyt et al. (2021) (there ±1 mag) to pass a set threshold as a quality indicator of the CMD, though is a smaller range to reduce the impact of photometric noise.

2. *# Stars below the tip, $N_{+,1.0}$*: The number of stars below the tip (applicability of Sobel filter). Following past studies, we calculate this as the total number of stars 1 magnitude below a tip such that

$$N_{+,1.0} = \text{number of stars 1 mag below tip.} \quad (9)$$

Madore et al. (2008) suggest that an effective TRGB measurement requires around 400 ∼ 500 stars within 1 mag below the tip, whereas Makarov et al. (2006) state that 50–100 is sufficient for a steady detection, though this is for the maximum-likelihood method. We include $N_{+,1.0} = 50$, 100, 200, and 400 as lower range limits for this optimization.

When we apply quality cuts on $[\mathrm{EDR_{tot}}]$, we create a new array of possible peaks, which we call $[\mathrm{EDR_{qual}}]$.

In order to further remove spurious peaks, we find the maximum peak in EDR space

$$\mathrm{EDR_{Max}} = \max([\mathrm{EDR_{qual}}]). \quad (10)$$

After that, we qualify every peak in $[\mathrm{EDR_{tot}}]$ (no matter its value of R) if it has a peak value greater than $0.6 \times \mathrm{EDR_{Max}}$, and we denote the set as $[\mathrm{EDR_{Peak}}]$. A threshold scale of 0.6 is a conservative value chosen to keep the real tip included, since it may not be located at the maximum derivative location (i.e., maximum EDR). The intention of this algorithm is to find all the possible peaks, retain most of them by applying fairly loose initial selection criteria, and plan to further refine the selection in the rest of the paper.

We correct the tip magnitude for both MW extinction (Schlafly & Finkbeiner 2011) and predicted host-galaxy extinction based on the distance to the center of the galaxy and the galaxy shape (Ménard et al. 2010). These corrections are on the small (on the ∼0.01 mag level) and discussed further in Section 5.1.

4. Results

4.1. Metrics for Tip Detection

Ultimately, we are attempting to minimize the dispersion of tip measurements between fields, while maintaining as many fields with a detection as possible. To evaluate the impact of different selection criteria, including spatial clipping, color band selection, smoothing scale, and R and $N_{+,1.0}$ filter, we introduce three different metrics:

1. σ_{tot}. The dispersion of magnitudes of all valid peaks that pass certain selection criteria. For a global σ_{tot} value for multiple galaxies, we weight the σ_{tot} for each galaxy by the number of tip measurements within that galaxy.
2. P_{valid}. The percentage of fields with a peak that passes the criteria. The ideal scenario is $P_{\mathrm{valid}} = 100\%$.
3. N_{TPF}. The average number of tips detected per field (TPF). This is calculated by dividing the number of valid peaks that pass selection criteria by the number of fields with at least one valid peak. The ideal scenario is $N_{\mathrm{TPF}} = 1$ as a value >1 implies some ambiguity in the identification of the peak.

In the Appendix, we discuss our optimization process of the values to be used from Table 1, and only the optimized parameters are used in the analysis unless otherwise stated. The optimization uses data from all the fields in galaxies with multiple fields.

4.2. A Case Study of TRGB Detection with NGC 3031

We focus on NGC 3031 as a *training set* to establish a robust algorithm and then we will apply it to other galaxies. In Figure 5, we show the EDR for a nominal analysis case with Hatt style weighting and both spatial and color clipping applied as previously described in Sections 3.1 and 3.2. We show the

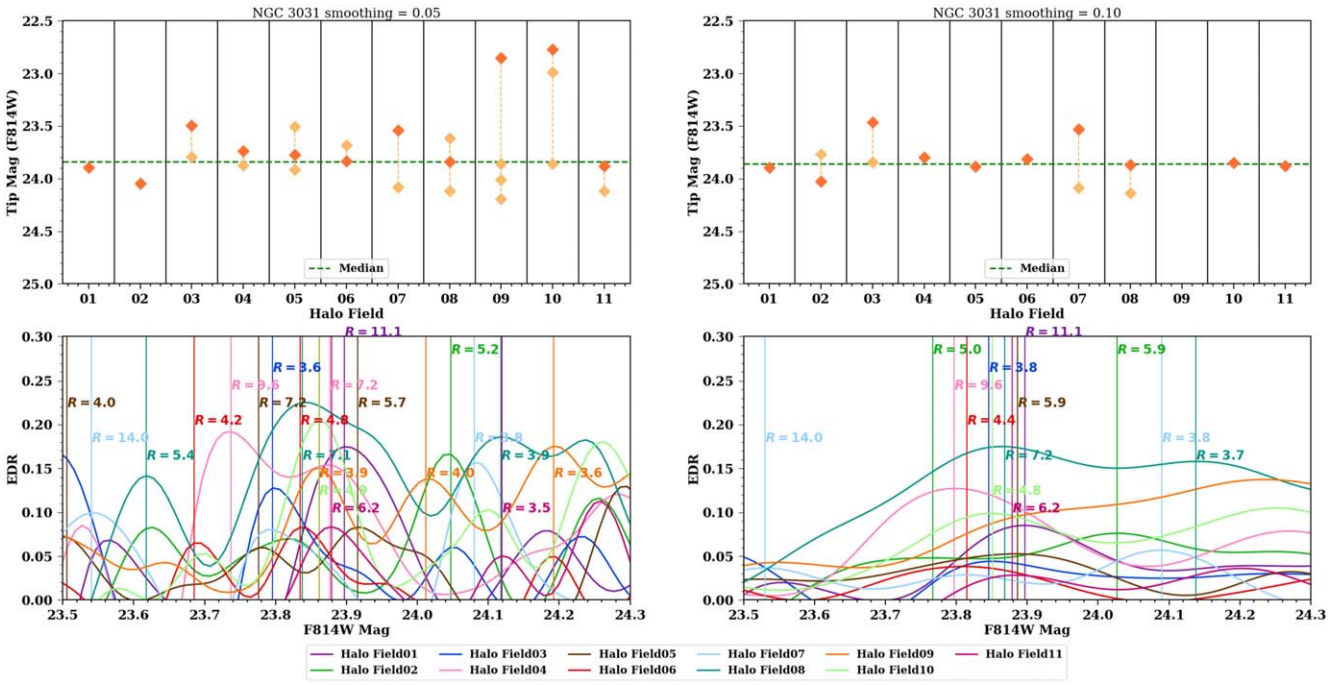

Figure 5. The tip magnitudes (top) and EDR (bottom) are shown for all the non-disk fields of NGC 3031. Note that the F814W magnitude ranges are different between the top panels and bottom panels, and the lower panels are zoomed in to show the wiggle patterns better and do not include the extreme outliers seen in the top panels. The left and right figures present the results for two different smoothing scales ($\sigma = 0.05$ and 0.1). In the top panels, the green lines denote the median of detections. We apply our algorithm to the individual field, labeling each tip detection with an orange diamond marker in the top panels, and the one with the highest R for each field is shaded darker, where R is the contrast defined in Equation (8). As shown in the bottom panels, there is little correlated structure of EDRs, except at the location of the best-fit line around F814W $\sim$ 23.85 mag in the bottom-right panel. Furthermore, for multiple fields, there are multiple peaks in this space, and an optimal algorithm must attempt to distinguish between these. Sometimes, the highest EDR peak is not aligned with the median as it is a false detection due to large absolute LF values at the faint end, and Equation (6) cannot sufficiently suppress that. We find that R is a critical indicator to characterize the quality of tip detections, which can filter out false tip detections fainter than the real tip. In this figure, only peaks with $R > 3.5$ are marked (top panel) and labeled (bottom panel, by the vertical line) as plausible. To further constrain the tip location, we need to apply $N_{+,1.0}$, which will filter out false detections in the bright end due to a small number of statistics, see Figure 6.

responses for two different smoothing values ($\sigma = 0.05$, 0.10) for the LF. We can see the dissimilarity of the shape and wiggles of the edge responses, even with the high S/N data, revealing true stochasticity likely from astrophysical variations in the field CMDs. As can be seen from the bottom panels in Figure 5, $\sigma = 0.10$ merges most adjacent peaks found for $\sigma = 0.05$ into a single peak. Further, we find the R selection can filter out false detections at the faint end, and the $N_{+,1.0}$ selection can filter out the extremely large R tips at the bright end (likely due to AGB stars). This can be seen from the shape of LF in Figure 3, whose value is smaller at the bright end. Consequently, $N_{+,1.0}$ selection will filter out detections at the bright end where star density is small. The full detection files and how R and $N_{+,1.0}$ make selections can be found on our website.[8]

A table that summarizes the results of the analysis steps discussed in Section 3 is shown in Table 2. The steps include spatial clipping, color band filtering, ratio R cut, $N_{+,1.0}$ cut, and smoothing scale modification. We apply these methods sequentially and cumulatively to show the improvement in the recovered dispersion. We find for an initial (Step 1 in Table 2) assessment of tip detection without any of the optimization choices described in the previous discussion, the total dispersion is $\sigma_{\rm tot} = 1.34$ and there is a mean of approximately nine detected tips per field indicating the wiggly nature of the EDR. The different steps have different impacts on either $\sigma_{\rm tot}$, $P_{\rm valid}$, or $N_{\rm TPF}$ defined in Section 4.1 as seen in Table 2. For instance, R and $N_{+,1.0}$ cuts have better efficiency in decreasing $\sigma_{\rm tot}$ and $N_{\rm TPF}$ than spatial clipping and color band. As can be seen in Figure 5, even after the spatial removal of young regions and the selection by color band, there is high dispersion due to the presence of multiple peaks in the Sobel response for each field. Spatial clipping and color band can only suppress the noise introduced by non-RGB stars, but cannot change the wiggling pattern of EDR much, thus slightly decreasing $\sigma_{\rm tot}$ and $N_{\rm TPF}$. On the contrary, R and $N_{+,1.0}$ provide a quality selection for the peaks of wiggling EDR, where R blocks peaks at the faint end, and $N_{+,1.0}$ denies peaks at the bright end. As shown in Figure 8, sufficiently large R can decrease $\sigma_{\rm tot}$ and $N_{\rm TPF}$ to extremely ideal values, regardless of the threshold value of $N_{+,1.0}$. Therefore, we find that the selection of tips with greater values of R is the most powerful method for eliminating false measurements while reducing the intra-host dispersion. Since there are still outliers in Figure 5 after a moderate $R > 3.5$ selection, $N_{+,1.0}$ is also significant to further confine the tip detections, especially the false detections around the edge of AGB. We also find that the $N_{+,1.0}$ cut reduces the dispersion but also reduces $P_{\rm valid}$ substantially, from 100% to below 50%. Moreover, we find a relatively larger smoothing scale is beneficial, as shown in Figure 5 and the Appendix because it can merge nearby peaks and suppress the wiggly EDR to give smaller $\sigma_{\rm tot}$ and $N_{\rm TPF}$.

Ultimately, we find we can reach $\sigma_{\rm tot} = 0.088$ mag for this host while $\sim$45% of the fields pass the entirety of selection

[8] https://github.com/JiaxiWu1018/Unsupervised-TRGB

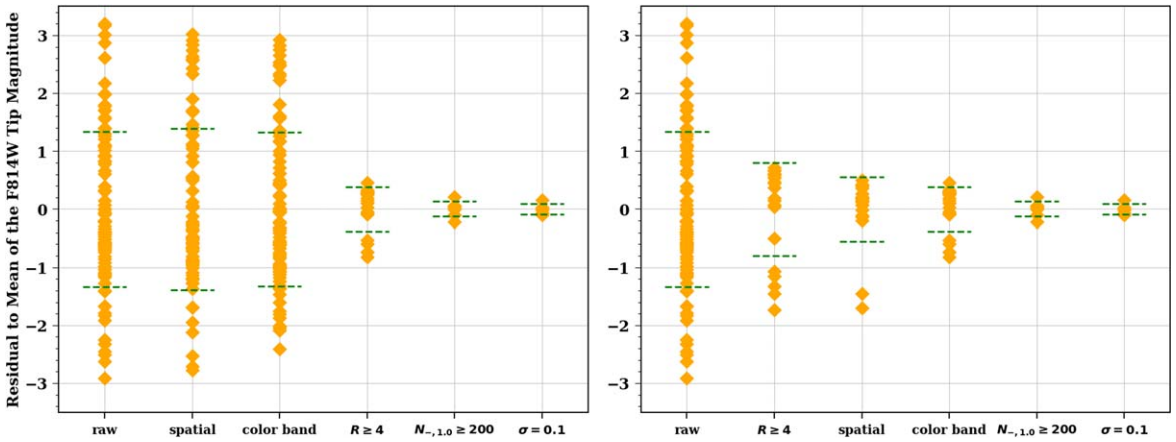

Figure 6. (Left) Following Table 2, for NGC 3031, tip residuals to the mean for all tip measurements. The residuals are shown sequentially for the different analysis steps. The range of green dashed lines shows the standard deviation of the TRGB detection set of each step. (Right) The same as the left, but with the R cut applied as the second cut. From both panels, we find that all analysis steps contribute to decreasing the overall dispersion, with the cut on the contrast ratio R as the strongest step.

Table 2
Quality of Tip Measurements of NGC 3031 and all Sample Fields with Different Analysis Steps Applied

Step	Description	Specification	Data	$\sigma_{\rm tot}$ (mag)	$P_{\rm valid}$	$N_{\rm TPF}$
1	Raw data	Equation (3): $\sigma = 0.05$, no other cuts	NGC 3031	1.34	100%	8.91
			ALL	1.35	100%	7.90
2	+Spatial clipping	Section 3.1, Figure 4: $\geqslant$10% max. blue star density	NGC 3031	1.39	100%	7.27
			ALL	1.30	100%	7.18
3	+Color band	Equation (2): $-7 < s < -5$, $w = 1$, free b	NGC 3031	1.32	100%	7.82
			ALL	1.21	100%	6.42
4	+R cut	Equation (8): $R \geqslant 4$	NGC 3031	0.38	100%	1.73
			ALL	0.47	94.00%	1.57
5	+ $N_{+,1.0}$ cut	Equation (9): $N_{+,1.0} \geqslant 200$	NGC 3031	0.13	45.45%	1.40
			ALL	0.080	56.00%	1.28
6	+ High smoothing	Equation (3): $\sigma = 0.1$	NGC 3031	0.088	45.45%	1.20
			ALL	0.058	52.00%	1.08

Notes. Columns represent, in turn, (1) step number; (2) description of the new technique applied in this step, with all techniques in former steps maintained if not contradictory (the only exception is in step 6, σ is changed to 0.1, not 0.05 in steps 1–5); (3) specific parameters used in this technique; (4) whether a data sample is NGC 3031 (11 fields total) or ALL galaxies (50 fields total); (5) dispersion of multi-field tip measurements around host mean; (6) percentage of fields with valid detection; (7) average number of tips per field (where >1 means a non-unique tip). As can be seen, $\sigma_{\rm tot}$ improves significantly with steps 4 and 5, but $P_{\rm valid}$ is reduced. We note a distinction in the description of the cuts on minimum R and $N_{+,1.0}$ that are applied here (the $R \geqslant 4$, and the $N_{+,1.0} \geqslant 200$) and the measured average values of R and $N_{+,1.0}$ of the qualified tips after application of these cuts (for this case $<R> \sim 7$ and $<N_{+,1.0}> \sim 2000$, respectively).

requirements, as shown in Step 6 of Table 2. For this case, we measure an $N_{\rm TPF} = 1.2$, still not quite one tip for each field, and the deviation from 1 is due to a single field with two tip detections. We show a graphical representation of the reduction in dispersion for NGC 3031 in Figure 6. We also show how the residuals are improved if we change the step order. As seen, the R cut is the most powerful one, since it decreases the dispersion from 1.32 to 0.38 and the $N_{\rm TPF}$ goes down from 7.82 to 1.73 for the first application sequence. Besides, we find that R pulls down the initial dispersion from 1.34 to 0.80 if applied directly to raw data. As a comparison, we calculate the impact of applying only $N_{+,1.0}$ cut to the initial data set of $\sigma_{\rm tot} = 1.34$, and we find $N_{+,1.0}$ can only decrease the total dispersion to 1.15. Therefore, R is the most powerful selection rule, though $N_{+,1.0}$ is necessary to further refine the validity of tip detections.

While we focus on the total dispersion of the recovered tip values, a related question is whether the location of the measured peaks changes due to the various analysis steps. In practice, this is difficult to quantify because there are often multiple tips detected for a specific field. However, if we attempt to trace the displacement of tip measurements due to the analysis steps, we find two general trends. The first is that

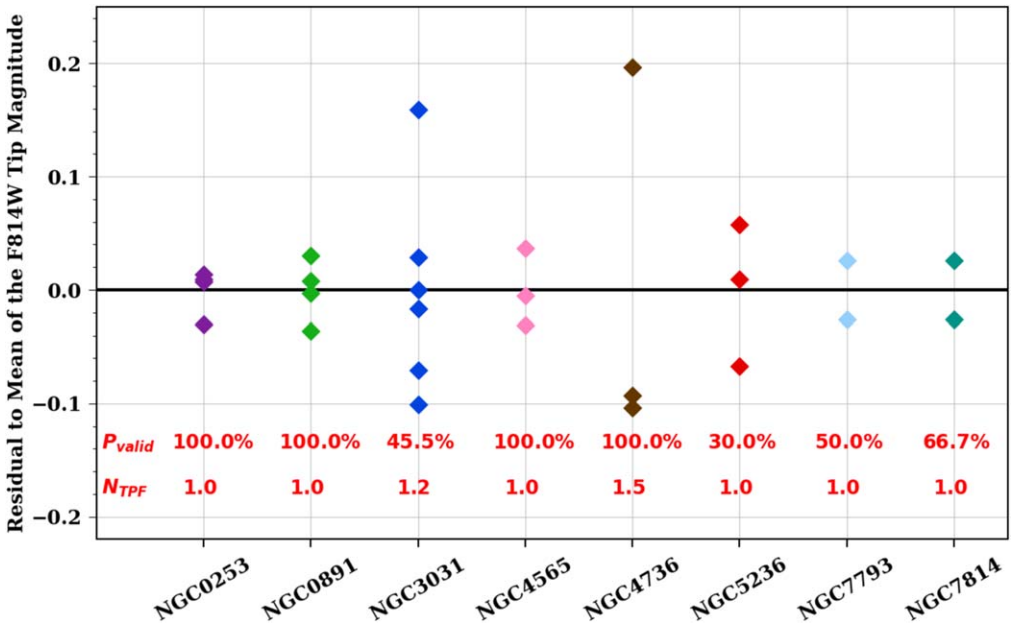

Figure 7. The dispersion of TRGB measurements from individual fields relative to the mean TRGB across fields for that galaxy. This is done for all 10 GHOSTS galaxies in which more than two fields are observed. A minimum R cut of 4 and $N_{+,1.0}$ cut of 200 are applied for this calculation (after which two of the hosts no longer have two valid fields).

the non-smoothing steps (Steps 2–5 in Table 2) have the impact of producing a brighter detected tip compared to that recovered in Step 1, typically on the order of ∼0.035 mag. We find the two steps that cause this shift are the spatial and color cuts. The smoothing step has more of a stochastic effect on the recovered tip magnitude, with impacts on the order of ±0.05 mag for a change in smoothing from 0.05 to 0.1.

When considering which fields will yield the best TRGB measurements, it is interesting to consider the example of the spatial positions of the fields in NGC 3031 and Figure 1. Top-quality measures ($R \geqslant 7$ and $N_{+,1.0} \geqslant 200$) are available only in three of the 11 Halo fields 01, 04, and 08. Fields 02 and 10 have only fair contrast ($R \sim 4$–6 with field 02 having two possible tips separated by ∼0.2 mag) and 05 and 11 have fair contrast and low $N_{+,1.0}$ (<200). The rest, fields 03, 06, 07, and 09 are of low quality and do not yield a reliable tip measurement. We find the best detection comes from the region just outside the disk (in this case just beyond the 25th mag arcsec^{-2} isophot and at a projected radius of 15–20 kpc). However, further fields are poorly constrained due to the small stellar population, which is hard for the Sobel filter to work well and the detections will be excluded by $N_{+,1.0}$ cut.

4.3. Dispersion of Fields from All GHOST Galaxies

With a processing algorithm shown to be successful with NGC 3031, we apply it to the nine other GHOSTS galaxies with multiple fields. We show the improvement in the peak measurements for all the galaxies in Table 2 and find results similar to what we found for NGC 3031. We find that $\sigma_{\rm tot}$ drops from 1.35 to 0.058 when including all the steps of the algorithm. The dispersion of tip brightnesses measured per galaxy can be seen in Figure 7. We find, similar to NGC 3031, that ∼50% of the fields pass these quality cuts. We also find that $N_{\rm TPF} = 1.08$ for the full set, showing that we measure close to one tip per field, which is the ideal scenario. There is one notable galaxy, i.e., NGC 4736 in Figure 7, which has one outlier that is 0.3 mag away from the other two detections. This is because, in one of the two fields of NGC 4736, there are two detections that pass through the selection criteria of $N \geqslant 4$, $N_{+,1.0} \geqslant 200$.

We show the key trends of our three metrics in Figure 8. For each of these metrics, we vary the minimum R ($R_{\rm min}$) value and calculate these metrics (left panel). We also show these trends as a function of the mean R using a rolling bin size of 3. We separate the measurements for fields of different minimum $N_{+,1.0}$ values at $N_{+,1.0} = 50$, 100, 200, and 400. We find a significant trend in the reduction of $\sigma_{\rm tot}$ with $R_{\rm min}$; this reduction is sharpest at $R_{\rm min} \sim 2.5$, but continues to the highest $R_{\rm min}$ values.

For high values of $R_{\rm mean} \sim 8$ and $N_{+,1.0} \geqslant 1000$, the dispersion reaches a very low ∼0.03 mag indicating the successful approach of the algorithm. For lower $N_{+,1.0}$ values and lower R values (<4), the dispersion may be as high as 0.8–0.9 mag owing to the frequent occurrence of multiple tips per field. We also show that $P_{\rm valid}$ steadily declines as $R_{\rm min}$ increases. For R values higher than 6, we are unable to recover a tip value for more than 50% of the fields. We also show that as we increase the quality requirements, $N_{\rm TPF}$ becomes closer to 1.

We can use the multi-field dispersion to approximate the uncertainty for a tip measurement as a function of its individual value of R and $N_{+,1.0}$ (instead of a group minimum threshold). We find that the multi-field dispersion (and hence, the TRGB uncertainty) can be roughly modeled by

$$\sigma = \sqrt{\left[\left(\frac{2e^{1.5(3-R)}}{e^{1.5(3-R)}+1}\right)\left(\frac{1}{N_{+,1.0}-100}\right)^{0.1}\right]^2 + 0.04^2} \text{ mag}, \quad (11)$$

where R and $N_{+,1.0}$ are the specific values for the field. This formula is roughly modeled from an ad hoc Fermi distribution, and parameters are set to best fit the numerical results. Also, note that this formula is only valid for $N_{+,1.0} > 100$. As the formula

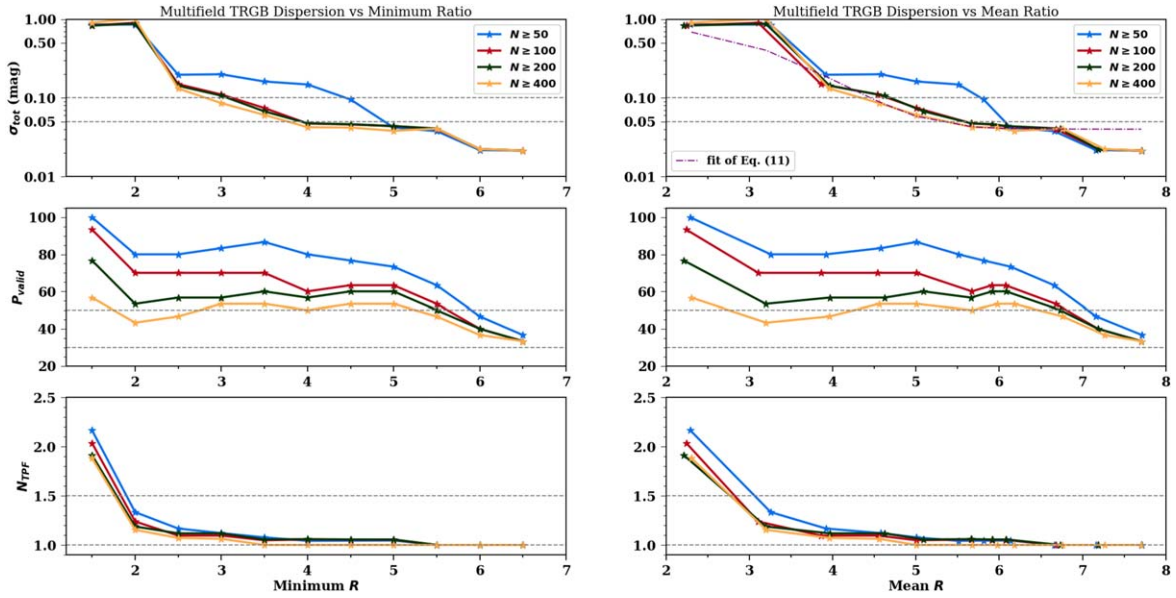

Figure 8. Values of the metrics defined in Section 4.1, as functions of different R and $N_{+,1.0}$, where R is the contrast ratio defined in Equation (8) and $N_{+,1.0}$ is the number of stars within one mag below tip as defined in Equation (9). In each panel, trend lines for different $N_{+,1.0}$ cuts are given. (Left) For each trend line, the metric is plotted with varying R cut (i.e., the minimum R as labeled on the x-axis), starting from 1.5–6.5. Note that for a specific R cut, the data which passes through this cut has a mean ratio typically 2–3 times greater than the cut limit. Likewise, the mean number of stars is a factor of a few greater than the minimum, see Section 4.3. (Left, top) The trend of the dispersion of field-tip magnitudes $\sigma_{\rm tot}$ as a function of the minimum ratio of R. (Left, middle) Similarly, the percentage of fields $P_{\rm valid}$ passes the cuts from the above panel. (Left, bottom) The number of detected tips per field $N_{\rm TPF}$ as a function of the minimum ratio of R. (Right) Changing the x-axis to be *mean* R instead of minimum, using a running bin size of 3 to make the set (e.g., for a minimum $R = 3$, we calculate the mean R of detections whose ratio $3 \leqslant R < 6$). (Right, top) Analogous to the left, but we add on one more curve which is the predicted $\sigma_{\rm tot}$ from Equation (11). (Right, middle/bottom) Analogous to the left.

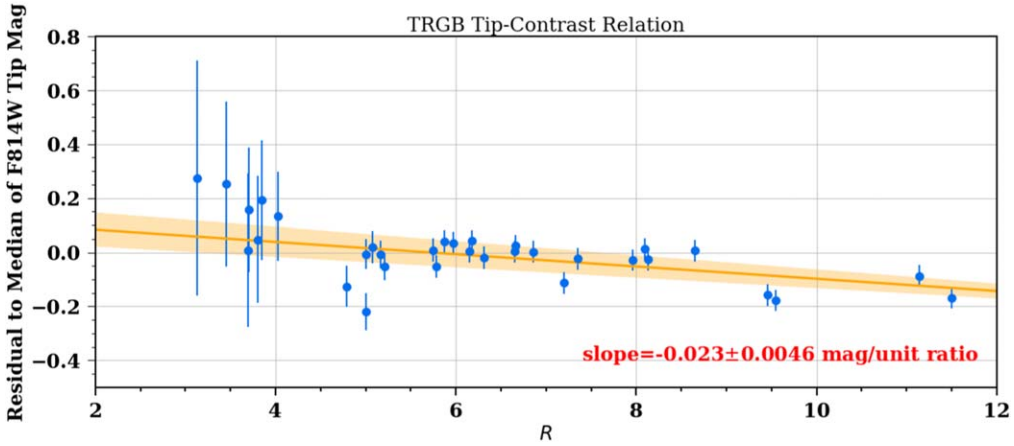

Figure 9. The relationship between measured tip-brightness residual for detected field TRGBs and R, using a $\sigma = 0.1$ smoothing and $R \geqslant 3$, $N_{+,1.0} \geqslant 200$ cut, where R is the contrast ratio defined in Equation (8), and $N_{+,1.0}$ is the number of stars within one mag below tip as defined in Equation (9). For each galaxy, we calculate each detection's residual to the median of all the TRGB measurements within that galaxy. Note that for each field of that specific galaxy, it is also likely to have multiple detections. The residuals are plotted vs. the detection's R, with error bars derived from Equation (11). The orange line is an error-weighted fitting of the blue data points, and the shaded region represents the error of this fitting. We find the slope of the orange fitting line to be -0.023 ± 0.0046 mag/unit ratio. Consequently, we find a TCR, which shows higher R is related to brighter tip magnitude.

indicates, an ideal (but rarely observed) sharp tip with $R \geqslant 10$ and $N_{+,1.0} \geqslant 1000$ can yield an uncertainty of $\sigma \sim 0.04$ mag. The more typical case of a less sharp tip measurement with $R \sim 4$ and a well-populated tip $N_{+,1.0} \sim 500$ yields $\sigma \sim 0.15$ mag. A very fuzzy tip with poor contrast, $R = 3$, produces much larger uncertainties with σ ranging from 0.3–1 mag depending on whether the tip is well-populated ($N_{+,1.0} = 1000$) or not ($N_{+,1.0} = 100$). Considering tips with $R < 3$ invariably allows multiple values and more information must be used to select among the tips.

4.4. The Tip-contrast Relation

Our analysis thus far has shown that R is especially powerful for predicting the recovered dispersion of peak magnitudes. In Figure 9, we compare this contrast value to the residuals of tip measurements relative to the median of all tips in the same host. We notice a clear trend such that the higher the R value, the brighter the apparent tip. Assuming that uncertainties follow Equation (11), we fit a line to the trend shown in Figure 9. We find a best-fit slope of -0.023 ± 0.0046 mag, which is highly significant at $\sim 5\sigma$. If we remove one outlier at

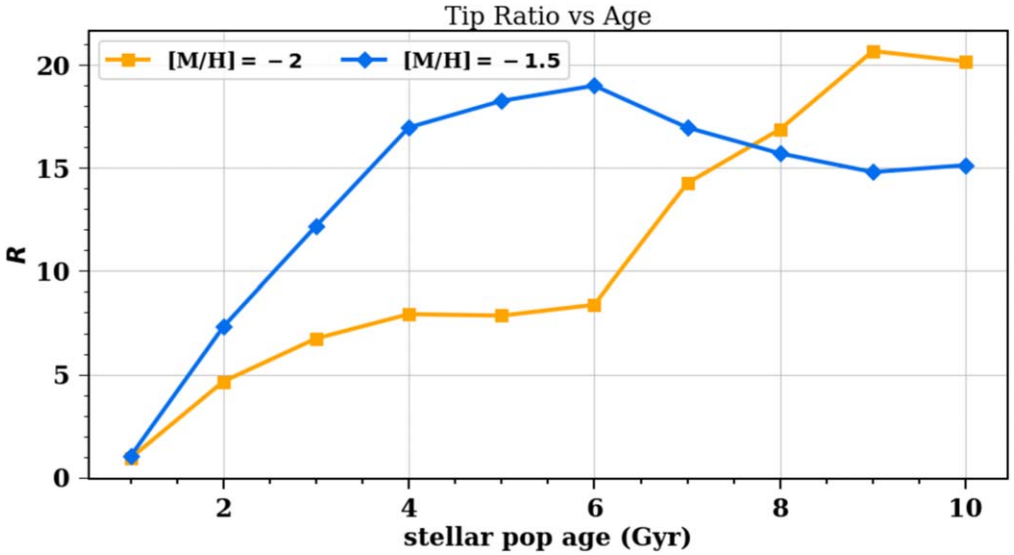

Figure 10. The Padova isochrones (version CMD3.7) were used to calculate LFs for the ACS F814W passband in the range of ±0.5 around the TRGB to calculate the contrast ratio. The calculations were done for single-age populations ranging from 1.5–10 Gyr at a low metallicity of [M/H] = −2 and −1.5. We plot the same RGB to AGB ratio R used throughout the empirical analysis vs. the age of the stellar population. As shown, the tip contrast increases with the age of the population, with a different dependence but the same general trend for different metallicities. The TCR relation found in Figure 9 may therefore show a relation between the inferred luminosity and the age of the stellar population.

$R \sim 5$ with a residual ~ -0.2 mag, which is $\sim 3.5\sigma$ away, the fit has a slope of -0.025 ± 0.0047. An unweighted fit gives a slope of -0.033 ± 0.0072 mag (4.5σ). If we further restrict $N_{+,1.0} \geqslant 400$, a weighted fit gives a slope of -0.029 ± 0.0054 mag, and an unweighted fit results in a slope of -0.037 ± 0.0071 mag (both with a 5σ significance). The host extinction, as discussed in Section 5.1, has little effect on this relation: if we increase the host extinction derived from Ménard et al. (2010) by an order of magnitude, the relation of the weighted fitting is little changed to -0.020 ± 0.0046 mag for $N_{+,1.0} \geqslant 200$, and -0.025 ± 0.0054 mag for $N_{+,1.0} \geqslant 400$. It also matches the scale and sense seen in other multi-field studies such as in NGC 4258 (Jang et al. 2021, see Figure 12) and in LMC fields (Hoyt et al. 2021).

Using this relation to standardize a set of tips of mixed contrast ratios (e.g., $R \geqslant 3$) lowers their dispersion by ∼10%–30% (depending on the number of stars). The standardization can be achieved by setting a fiducial point (we use $R = 4$) and correcting the tip magnitude counting in the tip-contrast relation (TCR) with a slope of −0.023 mag/unit ratio. For greater numbers of stars $N_{+,1.0} \geqslant 400$ (and a broad range of contrasts, $R \geqslant 3$) we see the greatest improvement with the dispersion declining from 0.095 to 0.074 mag.

5. Discussion and Conclusions

5.1. Impact of Dust Extinction

There are two sources of dust extinction that impact our results. The first is extinction due to the MW. Many of the fields are separated by angular separations that are comparable to the resolution of the MW dust map (Schlafly & Finkbeiner 2011) of ∼5′. We find on average variations of ∼3 mmag between fields, significantly smaller than the dispersion shown in Figure 6. A second source of dust extinction is internal to the host galaxy. However, because the fields are far from the center of the host, internal extinctions will be small. To quantify and account for these, we follow (Ménard et al. 2010, Equation (30)) to estimate the host extinction defined by the projected radii of the field from the core of a galaxy. The median field extinction estimate is 0.012 mag (standard deviation = 0.006 mag) and while included, has little impact on the TCR.

5.2. Understanding the TRGB TCR

In Figure 9, we presented a 5σ trend between tip-brightness residuals and measured R values. To better understand the origin of this relation, we use stellar population models to find the relation between stellar population age and contrast, as shown in Figure 10. These calculations were carried out for single-age populations with ages ranging from 1.5–10 Gyr at a metallicity of [M/H] = −2 and [M/H] = −1.5 using the Padova CMD3.7 stellar models from Girardi et al. (2010). As can be seen, there is a strong relation between age and R. This finding is consistent with a result in Girardi et al. (2010), which studies old metal-poor galaxies and found R values are all >20, significantly above any found in this work. From model predictions of McQuinn et al. (2019), it is expected that the F814W absolute luminosity of the TRGB should have a small dependence on stellar age of roughly 0.02 mag across an age range of 5 Gyr and 0.04 mag across 0.5 dex of metallicity. It is thus possible that the variation in R is related to population differences which cause an apparent R-luminosity relation at the level seen here.

It is also possible that the value of R correlates with patchy dust (i.e., anisotropic around a host as opposed to the mean, isotropic dust quantified by Ménard et al. (2010) and removed from the tips) with that dust as an underlying cause of the TCR. For the purpose of standardizing TRGB distance measurements, the cause of an apparent TCR is less material. For any of these causes, standardization and calibration using R would remove the present variation of the tip with R along the distance ladder and could also remove related biases due to differences in field demographics.

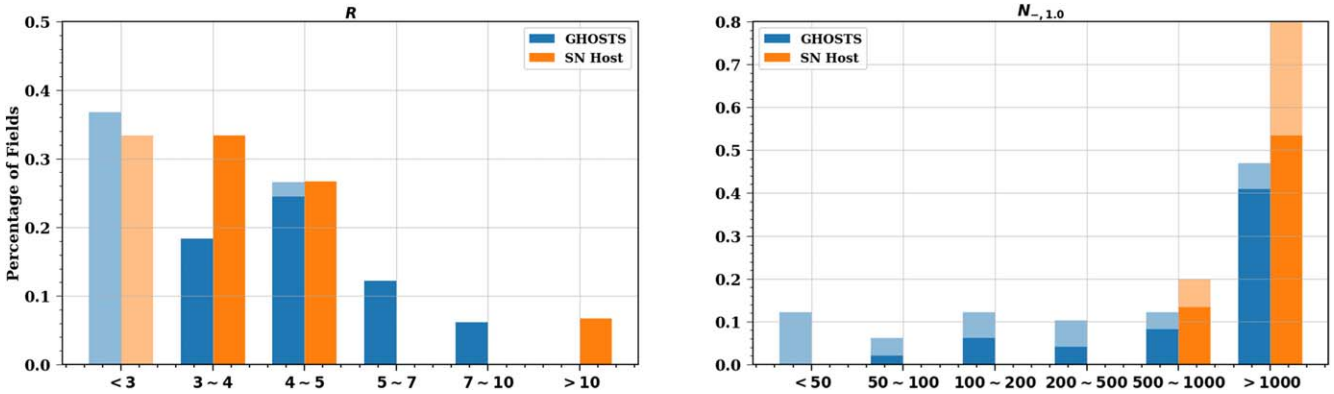

Figure 11. Histogram of the raw CMD properties discussed in Section 3.4 for the GHOSTS fields and the SN fields used in Anand et al. (2022). The lightly shaded blue and orange bars show the numbers for fields with either $R < 3$ or $N_{+,1.0} < 50$. We find generally good agreement between the distributions for the GHOSTS sample and SN hosts.

We also note that we do not find any significant evidence ($<1\sigma$) for a correlation between TRGB residuals and the mean color of the TRGB over the range of mean colors sampled by our tip measurements, 1.15 <F606W−F814W < 1.65 mag. Such a correlation may be present at our sensitivity level of $\leqslant$0.1 mag per magnitude change in color, but our use of a color band reduces our sensitivity to color differences.

5.3. Importance of the Hubble Constant

An important use of TRGB is in measurements of the Hubble constant (Freedman et al. 2019; Anand et al. 2022). In this role, the brightness of the TRGB in one or more *anchor* galaxies with geometric distance measurements is compared to that in *SNe host* galaxies. To quantify whether the inferences found for the GHOSTS set may be applied to the TRGB measurements in SN hosts, we compare the main diagnostics of these two sets, using the properties CMDs/TRGB catalog (Anand et al. 2021b) on the Extragalactic Distance Database[9] (Tully et al. 2009), which provides TRGB distances and the underlying HST photometry for of order 500 galaxies. We calculate the R and $N_{+,1.0}$ values for these fields before applying a spatial cut or a color band. As shown in Figure 11, we find an overall good overlap of the R parameter, our key diagnostic, indicating good applicability. There are some SN hosts with $R \leqslant 4$ which according to Figure 8 may not yield a reliable TRGB (though spatial clipping may be used to increase R). We do find that the number of stars in the SN fields is somewhat higher than that found for the GHOSTS fields, though this is likely due to the fact that the SN fields are more distant on average. Externally, we also examine other properties like RGB width and slope, and do not find any obvious inconsistencies.

While differences in tip magnitude across different fields in the same galaxy have been seen in many previous analyses (e.g., Hoyt et al. 2021; Jang et al. 2021; Anand et al. 2022), some of these analyses have chosen to favor brighter tips measured in some fields while discarding fainter tips in others. The value of this approach would depend on the goal of the analysis. If the goal is to measure the brightest apparent magnitude attained by the TRGB and nothing else, this approach is reasonable. If, however, the goal is to produce a comparison or calibration of populations with similar luminosity, our results suggest it is necessary to standardize the luminosity of the fields by accounting for their differing contrasts. Indeed, differences in TRGB contrast may play a role in the ∼0.1 mag range of recent calibrations of the absolute luminosity of the TRGB (see Blakeslee et al. 2021, Table 3 and Li et al. 2022, see Table 4 and Figure 6). We hope to explore such standardization in future work.

[9] edd.ifa.hawaii.edu

5.4. Summary

The goals of this analysis were to establish unsupervised measurements of the TRGB and quantify the field-to-field dispersion for a single galaxy. We find a dispersion as small as ∼0.03 mag, however, less than 50% of GHOSTS fields have a high enough contrast to produce some precise measure. We measure significant trends of the dispersion with the R contrast parameter and motivate the use of quality cuts on this value in the future. Furthermore, we find a trend between tip magnitude and the R value at $\sim 5\sigma$. The slope of this relation is a trend of up to 0.3 mag over the range of usable R values seen for the GHOSTS fields (from $R = 3$–12). We do not see any obvious differences between the GHOST fields analyzed here and the fields analyzed for distance ladder measurements. In follow-up work, we will apply the standardization methods and insights found here to refine measurements of H_0 from TRGB.

Acknowledgments

D.S. is supported by Department of Energy grant DE-SC0010007, the David and Lucile Packard Foundation, the Templeton Foundation and Sloan Foundation. We greatly appreciate the GHOSTS team and EDD team for making all of their data public.

This research has made use of NASA's Astrophysics Data System.

Facilities: MAST, HST:ACS. Some of the data presented in this paper were obtained from the Mikulski Archive for Space Telescopes (MAST) at the Space Telescope Science Institute. The specific observations analyzed can be accessed via doi:10.17909/T9NC7Z.

Software: Astropy (Robitaille et al. 2013, Price-Whelan et al. 2018), Matplotlib (Hunter 2007), NumPy (van der Walt et al. 2011; Harris et al. 2020).

Data Availability

We make great efforts to allow the community to reproduce and improve on this analysis. We include all of the code used to make the plots in this analysis. The data comes from the GHOSTS program and is accessible at https://archive.stsci.edu/prepds/ghosts/ghosts/survey.html. For ease, we also re-release this data, along with our codes and plots, at https://github.com/JiaxiWu1018/Unsupervised-TRGB.

Appendix
Optimization

Here we show the plots (Figure 12) optimizing the analysis numbers discussed in Section 3. Our optimization is done to find the best values for the analysis parameters as described in Table 1. This optimization is done with data from all GHOSTS galaxies that have two or more halo fields.

Due to computational limitations, we perform a uni-dimensional optimization where we fix all parameters to a single value except for one parameter and vary that one alone. This is done iteratively to allow for convergence in this multidimensional space. We show the optimization for the width and slope of the color band, smoothing value of the LF, and weighting methods as defined in Equations (4)–(6) in the four panels of Figure 12, respectively. All optimizations are done using our three metrics discussed in Section 4.1.

Ultimately, we find some sensitivity to the bandwidth, limited sensitivity to the band slope, strong sensitivity to the smoothing value, and limited sensitivity to the weighting algorithm. The impact of the smoothing value appears to plateau around $\sigma = 0.05$, while the number of fields continues to decrease. The decrease in the number of fields with a tip is due to the $R \geqslant 4$ cut, where R is the contrast ratio defined in Equation (8). As a different tip is found with different smoothing, the value at R at that tip changes as well. We note that the change in R is not due to smoothing itself as it is calculated directly from the number of stars on the CMD, but rather due to the location of the peak. Therefore, for the

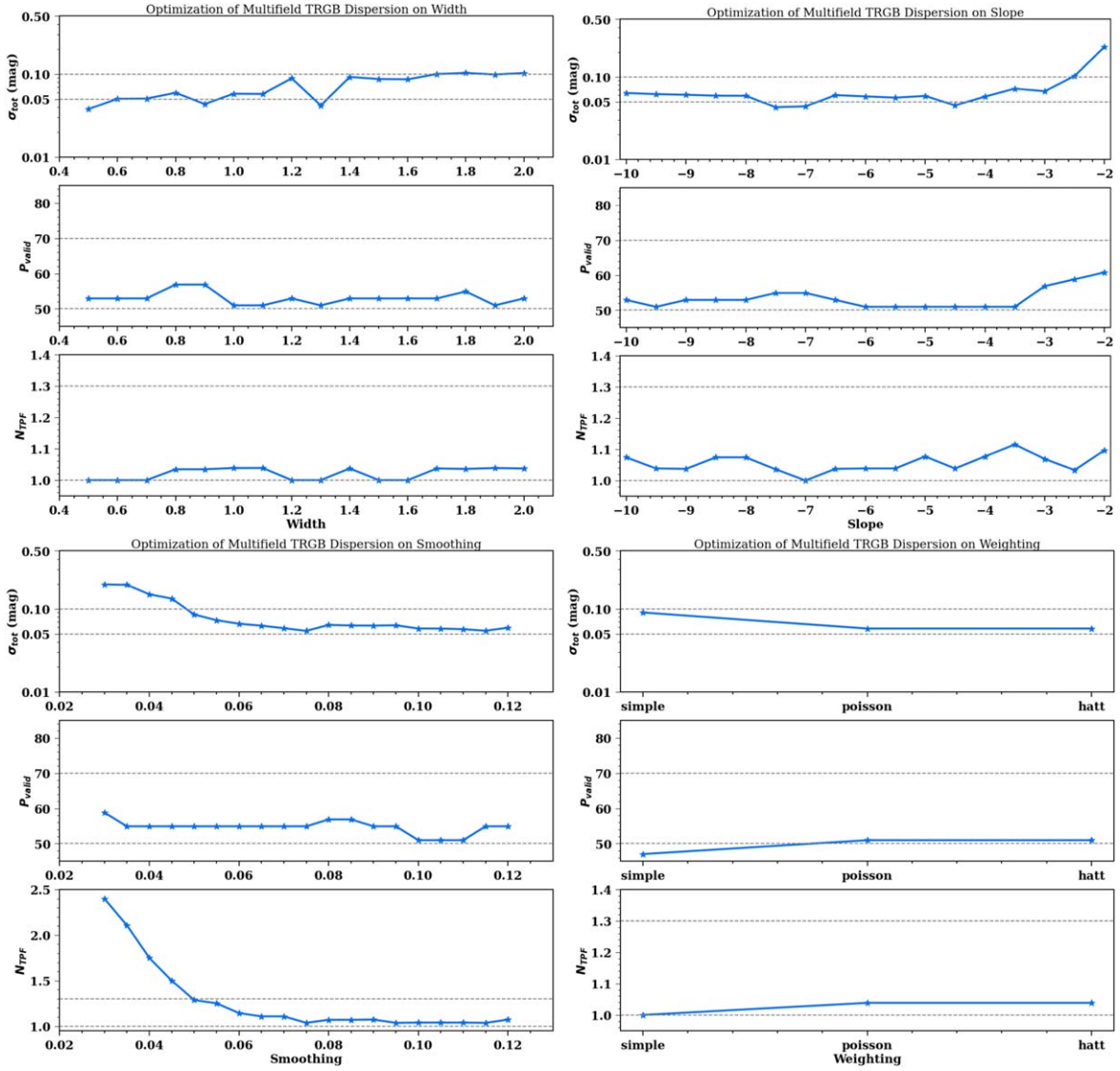

Figure 12. For the methods described in Table 2, the optimization is to determine the best values/choices for the different parameters. For each of our parameters (color bandwidth and slope, smoothing, weighting), we show trends of our three metrics when varying the values. The color bandwidth and slope are, respectively, w and s defined in Equation (2). The smoothing is discussed in Section 3.4 Equation (3), and the three weighting techniques are listed in Equations (4)–(6). The dashed horizontal lines are referential lines for better visual localization.

analyses, we select a color bandwidth of 1.0, and for GHOSTS CMDs which use F606W and F814W filters, we set the range of band slope to be [−7, −5]. For the smoothing of the LF, we use a scale of 0.10, and then provide the EDR of the smoothed LF using Equation (6).

ORCID iDs

J. Wu https://orcid.org/0000-0003-3829-967X
D. Scolnic https://orcid.org/0000-0002-4934-5849
A. G. Riess https://orcid.org/0000-0002-6124-1196
G. S. Anand https://orcid.org/0000-0002-5259-2314
R. Beaton https://orcid.org/0000-0002-1691-8217
S. Li https://orcid.org/0000-0002-8623-1082

References

Alam, S., Albareti, F. D., Allende Prieto, C., et al. 2015, ApJS, 219, 12
Anand, G. S., Lee, J. C., Van Dyk, S. D., et al. 2021a, MNRAS, 501, 3621
Anand, G. S., Rizzi, L., & Tully, R. B. 2018, AJ, 156, 105
Anand, G. S., Rizzi, L., Tully, R. B., et al. 2021b, AJ, 162, 80
Anand, G. S., Tully, R. B., Rizzi, L., Shaya, E. J., & Karachentsev, I. D. 2019, ApJ, 880, 52
Anand, G. S., Tully, R. B., Rizzi, L., Riess, A. G., & Yuan, W. 2022, ApJ, 932, 15
Beaton, R. L., Seibert, M., Hatt, D., et al. 2019, ApJ, 885, 141
Blakeslee, J. P., Jensen, J. B., Ma, C.-P., Milne, P. A., & Greene, J. E. 2021, ApJ, 911, 65
Cohen, R. E., Goudfrooij, P., Correnti, M., et al. 2020, ApJ, 890, 52
Condon, J. J. 1987, ApJS, 65, 485
Condon, J. J., Helou, G., Sanders, D. B., & Soifer, B. T. 1996, ApJS, 103, 81
Conn, A. R., Ibata, R. A., Lewis, G. F., et al. 2012, ApJ, 758, 11
Crnojević, D., Sand, D. J., Bennet, P., et al. 2019, ApJ, 872, 80
Dhawan, S., Goobar, A., Johansson, J., et al. 2022, ApJ, 934, 185
Dolphin, A. E. 2000, PASP, 112, 1383
Freedman, W. L., Madore, B. F., Hatt, D., et al. 2019, ApJ, 882, 34
Freedman, W. L., Madore, B. F., Hoyt, T., et al. 2020, ApJ, 891, 57
Girardi, L., Williams, B. F., Gilbert, K. M., et al. 2010, ApJ, 724, 1030
Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, Natur, 585, 357
Hatt, D., Beaton, R. L., Freedman, W. L., et al. 2017, ApJ, 845, 146
Hatt, D., Freedman, W. L., Madore, B. F., et al. 2018, ApJ, 866, 145
Heald, G., Józsa, G., Serra, P., et al. 2011, A&A, 526, A118
Hoyt, T. J. 2021, arXiv:2106.13337
Hoyt, T. J., Beaton, R. L., Freedman, W. L., et al. 2021, ApJ, 915, 34
Hoyt, T. J., Freedman, W. L., Madore, B. F., et al. 2019, ApJ, 882, 150
Hunter, J. D. 2007, CSE, 9, 90
Jang, I. S., & Lee, M. G. 2017, ApJ, 836, 74
Jang, I. S., Hatt, D., Beaton, R. L., et al. 2018, ApJ, 852, 60
Jang, I. S., Hoyt, T. J., Beaton, R. L., et al. 2021, ApJ, 906, 125
Lee, M. G., Freedman, W. L., & Madore, B. F. 1993, ApJ, 417, 553
Li, S., Casertano, S., & Riess, A. G. 2022, ApJ, 939, 96
Madore, B. F., & Freedman, W. L. 1995, AJ, 109, 1645
Madore, B. F., Mager, V., & Freedman, W. L. 2008, ApJ, 690, 389
Mager, V. A., & Hoyt, T. J. 2020, AAS Meeting Abstracts, 235, 279.10
Makarov, D., Makarova, L., Rizzi, L., et al. 2006, AJ, 132, 2729
McQuinn, K. B. W., Boyer, M., Skillman, E. D., & Dolphin, A. E. 2019, ApJ, 880, 63
McQuinn, K. B. W., Cannon, J. M., Dolphin, A. E., et al. 2014, ApJ, 785, 3
Ménard, B., Scranton, R., Fukugita, M., & Richards, G. 2010, MNRAS, 405, 1025
Méndez, B., Davis, M., Moustakas, J., et al. 2002, AJ, 124, 213
Persson, S. E., Madore, B. F., Krzeminski, W., et al. 2004, AJ, 128, 2239
Price-Whelan, A. M., Sipőcz, B. M., Günther, H. M., et al. 2018, AJ, 156, 123
Radburn-Smith, D. J., de Jong, R. S., Seth, A. C., et al. 2011, ApJS, 195, 18
Robitaille, T. P., Tollerud, E. J., Greenfield, P., et al. 2013, A&A, 558, A33
Schlafly, E. F., & Finkbeiner, D. P. 2011, ApJ, 737, 103
Scolnic, D., Smith, M., Massiah, A., et al. 2020, ApJL, 896, L13
Serenelli, A., Weiss, A., Cassisi, S., Salaris, M., & Pietrinferni, A. 2017, A&A, 606, A33
Shaya, E. J., Tully, R. B., Pomarède, D., & Peel, A. 2020, ApJ, 927, 168
Shen, Z., Danieli, S., van Dokkum, P., et al. 2021, ApJL, 914, L12
Tully, R. B., Kourkchi, E., Courtois, H. M., et al. 2023, ApJ, 944, 94
Tully, R. B., Rizzi, L., Shaya, E. J., et al. 2009, AJ, 138, 323
van der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, CSE, 13, 22
Walter, F., Brinks, E., de Blok, W. J. G., et al. 2008, AJ, 136, 2563